SEQUENCE LISTING

<110> Lim, Moon Young
      Edwards, Cynthia A.
      Fry, Kirk E.
      Bruice, Thomas W.
      Starr, Douglas B.
      Laurance, Megan E.
      Kwok, Yan

<120> DNA Binding Compound-Mediated Molecular
   Switch System

<130> 4600-0130.30

<140> US 09/518,297
<141> 2000-03-03

<150> US 60/122,513
<151> 1999-03-03

<150> US 60/154,605
<151> 1999-09-17

<160> 62

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA response element

<400> 1
cgttcgcact t   11

<210> 2
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA response element

<400> 2
cggagtactg tcctccg   17

<210> 3
<211> 12
<212> DNA
<213> Artificial Sequence

1

<220>
<223> DNA response element

<221> misc_feature
<222> (1)...(12)
<223> n = A,T,C or G

<400> 3
taattanggg ng   12

<210> 4
<211> 551
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (0)...(0)
<223> transcriptional regulatory protein

<400> 4

```
Met Asp Glu Leu Phe Pro Leu Ile Phe Pro Ala Glu Pro Ala Gln Ala
 1               5                  10                  15
Ser Gly Pro Tyr Val Glu Ile Ile Glu Gln Pro Lys Gln Arg Gly Met
                20                  25                  30
Arg Phe Arg Tyr Lys Cys Glu Gly Arg Ser Ala Gly Ser Ile Pro Gly
            35                  40                  45
Glu Arg Ser Thr Asp Thr Thr Lys Thr His Pro Thr Ile Lys Ile Asn
        50                  55                  60
Gly Tyr Thr Gly Pro Gly Thr Val Arg Ile Ser Leu Val Thr Lys Asp
65                  70                  75                  80
Pro Pro His Arg Pro His Pro His Glu Leu Val Gly Lys Asp Cys Arg
                85                  90                  95
Asp Gly Phe Tyr Glu Ala Glu Leu Cys Pro Asp Arg Cys Ile His Ser
            100                 105                 110
Phe Gln Asn Leu Gly Ile Gln Cys Val Lys Lys Arg Asp Leu Glu Gln
        115                 120                 125
Ala Ile Ser Gln Arg Ile Gln Thr Asn Asn Asn Pro Phe Gln Val Pro
    130                 135                 140
Ile Glu Glu Gln Arg Gly Asp Tyr Asp Leu Asn Ala Val Arg Leu Cys
145                 150                 155                 160
Phe Gln Val Thr Val Arg Asp Pro Ser Gly Arg Pro Leu Arg Leu Pro
                165                 170                 175
Pro Val Leu Pro His Pro Ile Phe Asp Asn Arg Ala Pro Asn Thr Ala
                180                 185                 190
Glu Leu Lys Ile Cys Arg Val Asn Arg Asn Ser Gly Ser Cys Leu Gly
            195                 200                 205
Gly Asp Glu Ile Phe Leu Leu Cys Asp Lys Val Gln Lys Glu Asp Ile
    210                 215                 220
Glu Val Tyr Phe Thr Gly Pro Gly Trp Glu Ala Arg Gly Ser Phe Ser
225                 230                 235                 240
Gln Ala Asp Val His Arg Gln Val Ala Ile Val Phe Arg Thr Pro Pro
                245                 250                 255
Tyr Ala Asp Pro Ser Leu Gln Ala Pro Val Arg Val Ser Met Gln Leu
                260                 265                 270
Arg Arg Pro Ser Asp Arg Glu Leu Ser Glu Pro Met Glu Phe Gln Tyr
            275                 280                 285
```

2

```
Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg
    290                 295                 300
Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly
305                 310                 315                 320
Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg
                325                 330                 335
Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr
                340                 345                 350
Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe
                355                 360                 365
Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro
    370                 375                 380
Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val
385                 390                 395                 400
Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly
                405                 410                 415
Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly
                420                 425                 430
Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu
                435                 440                 445
Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr
    450                 455                 460
Asp Leu Ala Ser Val Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln
465                 470                 475                 480
Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr
                485                 490                 495
Pro Glu Ala Ile Thr Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp
                500                 505                 510
Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu
                515                 520                 525
Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala
    530                 535                 540
Leu Leu Ser Gln Ile Ser Ser
545                 550
```

<210> 5
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA response element

<400> 5
tccctatcag tgatagaga   19

<210> 6
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> response element

<400> 6
cttaacactc gcgagtgtta ag   22

<210> 7
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> response element

<221> misc_feature
<222> (3)...(3)
<223> n = G or T

<221> misc_feature
<222> (7)...(7)
<223> n = A,T,C or G

<221> misc_feature
<222> (12)...(12)
<223> n = A or C

<400> 7
rgntcantga cny   13

<210> 8
<211> 77
<212> PRT
<213> Artificial Sequence

<220>
<223> activator sequence

<400> 8
Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp Gly
 1   5   10   15
Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp Leu
            20   25   30
Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro His
            35   40   45
Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe Glu
    50   55   60
Gln Met Phe Thr Asp Ala Leu Gly Ile Asp Glu Tyr Gly
65   70   75

<210> 9
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> activator sequence

<221> VARIANT
<222> (1)...(11)
<223> tetramer

<400> 9
Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu

4

<210> 10
<211> 97
<212> PRT
<213> Artificial Sequence

<220>
<223> repressor sequence

<400> 10
Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe
 1   5   10   15
Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp
            20   25   30
Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys
        35   40   45
Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu
    50   55   60
Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Glu Arg Glu Ile His
65   70   75   80
Gln Glu Thr His Pro Asp Ser Glu Thr Ala Phe Glu Ile Lys Ser Ser
                85   90   95
Val

<210> 11
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> repressor sequence

<400> 11
Met Ala Ala Ala Val Arg Met Asn Ile Gln Met Leu Leu Glu Ala Ala
 1   5   10   15
Asp Tyr Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser
            20   25   30
Met Leu Pro Tyr
        35

<210> 12
<211> 116
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (0)...(0)
<223> partial promoter sequence

<400> 12
cgcggtcaga aaattatttt aaatttcctc ttgtcaggcc ggaataactc cctataatgc   60
gccaccactg acacggaaca acggcaaaca cgccgccggg tcagcggggt tctcct   116

<210> 13

```
<211> 22
<212> DNA
<213> Escherichia coli

<220>
<221> misc_feature
<222> (0)...(0)
<223> partial promoter sequence

<400> 13
agaaaattat tttaaatttc ct                                              22

<210> 14
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> modified promoter sequence

<400> 14
gactgcagtg gtacctagga gg                                              22

<210> 15
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> modified promoter sequence

<400> 15
agaaaattat tttaaatttc ct                                              22

<210> 16
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> modified promoter sequence

<400> 16
ggaaaatttt ttttcaaaag ta                                              22

<210> 17
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> modified promoter sequence

<400> 17
tgaaatttat tttgcgaaag gg                                              22

<210> 18
```

<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 18
tgttcgcact t   11

<210> 19
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 19
catggacgcc actgagccgt ttttgttcgc acttgaggcg agtcgatgca cc   52

<210> 20
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 20
catggacgcc actgagccgt gttcgcactt tttttgagg cgagtcgatg cacc   54

<210> 21
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 21
catggacgcc actgagccgt ttttgttcgc acttttttt gaggcgagtc gatgcacc   58

<210> 22
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 22
cttaaaaata ac   12

<210> 23
<211> 16
<212> DNA

<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 23
ttgaaaaatc aacgct   16

<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 24
tttttgttcg cacttttttt t   21

<210> 25
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 25
tttttgggat tttccttttt   20

<210> 26
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 26
aaaaaattgt gagcgctcac aatttttt   28

<210> 27
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> tissue-specific transcription factor

<400> 27
acttta   6

<210> 28
<211> 9
<212> DNA
<213> Artificial Sequence

8

```
<220>
<223> engineered DNA response element

<400> 28
taccgacat                                                                          9

<210> 29
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 29
gggactttcc                                                                        10

<210> 30
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 30
gggattttcc                                                                        10

<210> 31
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 31
cgaccgtgct cgagttaacg ggactttcca aaaacgatcg gactggactc                            50

<210> 32
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 32
cgaccgtgct cgagttaacg ggattttcca aaaacgatcg gactggactc                            50

<210> 33
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element
```

<400> 33
cgaccgtgct cgagaaattg ggattttcca aaaacgatcg gactggactc   50

<210> 34
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 34
aaaaaattgt gagcgctcac aatttttt   28

<210> 35
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 35
ttttttttgt gagcggataa caaaa   25

<210> 36
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 36
tctgggatcc   10

<210> 37
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 37
gagttttttt taag   14

<210> 38
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered DNA response element

<400> 38

gagtttttaaa agag

<210> 39
<211> 969
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (0)...(0)
<223> transcriptional regulatory protein

<400> 39
```
Met Ala Glu Asp Asp Pro Tyr Leu Gly Arg Pro Glu Gln Met Phe His
 1               5                  10                  15
Leu Asp Pro Ser Leu Thr His Thr Ile Phe Asn Pro Glu Val Phe Gln
             20                  25                  30
Pro Gln Met Ala Leu Pro Thr Ala Asp Gly Pro Tyr Leu Gln Ile Leu
         35                  40                  45
Glu Gln Pro Lys Gln Arg Gly Phe Arg Phe Arg Tyr Val Cys Glu Gly
     50                  55                  60
Pro Ser His Gly Gly Leu Pro Gly Ala Ser Ser Glu Lys Asn Lys Lys
65                  70                  75                  80
Ser Tyr Pro Gln Val Lys Ile Cys Asn Tyr Val Gly Pro Ala Lys Val
                 85                  90                  95
Ile Val Gln Leu Val Thr Asn Gly Lys Asn Ile His Leu His Ala His
             100                 105                 110
Ser Leu Val Gly Lys His Cys Glu Asp Gly Ile Cys Thr Val Thr Ala
             115                 120                 125
Gly Pro Lys Asp Met Val Val Gly Phe Ala Asn Leu Gly Ile Leu His
         130                 135                 140
Val Thr Lys Lys Lys Val Phe Glu Thr Leu Glu Ala Arg Met Thr Glu
145                 150                 155                 160
Ala Cys Ile Arg Gly Tyr Asn Pro Gly Leu Leu Val His Pro Asp Leu
                 165                 170                 175
Ala Tyr Leu Gln Ala Glu Gly Gly Gly Asp Arg Gln Leu Gly Asp Arg
             180                 185                 190
Glu Lys Glu Leu Ile Arg Gln Ala Ala Leu Gln Gln Thr Lys Glu Met
             195                 200                 205
Asp Leu Ser Val Val Arg Leu Met Phe Thr Ala Phe Leu Pro Asp Ser
             210                 215                 220
Thr Gly Ser Phe Thr Arg Arg Leu Glu Pro Val Val Ser Asp Ala Ile
225                 230                 235                 240
Tyr Asp Ser Lys Ala Pro Asn Ala Ser Asn Leu Lys Ile Val Arg Met
                 245                 250                 255
Asp Arg Thr Ala Gly Cys Val Thr Gly Gly Glu Glu Ile Tyr Leu Leu
             260                 265                 270
Cys Asp Lys Val Gln Lys Asp Asp Ile Gln Ile Arg Phe Tyr Glu Glu
             275                 280                 285
Glu Glu Asn Gly Gly Val Trp Glu Gly Phe Gly Asp Phe Ser Pro Thr
         290                 295                 300
Asp Val His Arg Gln Phe Ala Ile Val Phe Lys Thr Pro Lys Tyr Lys
305                 310                 315                 320
Asp Ile Asn Ile Thr Lys Pro Ala Ser Val Phe Val Gln Leu Arg Arg
                 325                 330                 335
Lys Ser Asp Leu Glu Thr Ser Glu Pro Lys Pro Phe Leu Tyr Tyr Pro
             340                 345                 350
```

```
Glu Ile Lys Asp Lys Glu Glu Val Gln Arg Lys Arg Gln Lys Leu Met
    355                     360                 365
Pro Asn Phe Ser Asp Ser Phe Gly Gly Gly Ser Gly Ala Gly Ala Gly
    370                 375                 380
Gly Gly Gly Met Phe Gly Ser Gly Gly Gly Gly Gly Gly Thr Gly Ser
385                 390                 395                 400
Thr Gly Pro Gly Tyr Ser Phe Pro His Tyr Gly Phe Pro Thr Tyr Gly
                405                 410                 415
Gly Ile Thr Phe His Pro Gly Thr Thr Lys Ser Asn Ala Gly Met Lys
                420                 425                 430
His Gly Thr Met Asp Thr Glu Ser Lys Lys Asp Pro Glu Gly Cys Asp
    435                 440                 445
Lys Ser Asp Asp Lys Asn Thr Val Asn Leu Phe Gly Lys Val Ile Glu
    450                 455                 460
Thr Thr Glu Gln Asp Gln Glu Pro Ser Glu Ala Thr Val Gly Asn Gly
465                 470                 475                 480
Glu Val Thr Leu Thr Tyr Ala Thr Gly Thr Lys Glu Glu Ser Ala Gly
                485                 490                 495
Val Gln Asp Asn Leu Phe Leu Glu Lys Ala Met Gln Leu Ala Lys Arg
                500                 505                 510
His Ala Asn Ala Leu Phe Asp Tyr Ala Val Thr Gly Asp Val Lys Met
    515                 520                 525
Leu Leu Ala Val Gln Arg His Leu Thr Ala Val Gln Asp Glu Asn Gly
    530                 535                 540
Asp Ser Val Leu His Leu Ala Ile Ile His Leu His Ser Gln Leu Val
545                 550                 555                 560
Arg Asp Leu Leu Glu Val Thr Ser Gly Leu Ile Ser Asp Asp Ile Ile
                565                 570                 575
Asn Met Arg Asn Asp Leu Tyr Gln Thr Pro Leu His Leu Ala Val Ile
                580                 585                 590
Thr Lys Gln Glu Asp Val Val Glu Asp Leu Leu Arg Ala Gly Ala Asp
    595                 600                 605
Leu Ser Leu Leu Asp Arg Leu Gly Asn Ser Val Leu His Leu Ala Ala
    610                 615                 620
Lys Glu Gly His Asp Lys Val Leu Ser Ile Leu Leu Lys His Lys Lys
625                 630                 635                 640
Ala Ala Leu Leu Leu Asp His Pro Asn Gly Asp Gly Leu Asn Ala Ile
                645                 650                 655
His Leu Ala Met Met Ser Asn Ser Leu Pro Cys Leu Leu Leu Leu Val
                660                 665                 670
Ala Ala Gly Ala Asp Val Asn Ala Gln Glu Gln Lys Ser Gly Arg Thr
                675                 680                 685
Ala Leu His Leu Ala Val Glu His Asp Asn Ile Ser Leu Ala Gly Cys
    690                 695                 700
Leu Leu Leu Glu Gly Asp Ala His Val Asp Ser Thr Thr Tyr Asp Gly
705                 710                 715                 720
Thr Thr Pro Leu His Ile Ala Ala Gly Arg Gly Ser Thr Arg Leu Ala
                725                 730                 735
Ala Leu Leu Lys Ala Ala Gly Ala Asp Pro Leu Val Glu Asn Phe Glu
                740                 745                 750
Pro Leu Tyr Asp Leu Asp Asp Ser Trp Glu Asn Ala Gly Glu Asp Glu
    755                 760                 765
Gly Val Val Pro Gly Thr Thr Pro Leu Asp Met Ala Thr Ser Trp Gln
    770                 775                 780
Val Phe Asp Ile Leu Asn Gly Lys Pro Tyr Glu Pro Glu Phe Thr Ser
785                 790                 795                 800
Asp Asp Leu Leu Ala Gln Gly Asp Met Lys Gln Leu Ala Glu Asp Val
```

```
                    805                     810                     815
Lys Leu Gln Leu Tyr Lys Leu Leu Glu Ile Pro Asp Pro Asp Lys Asn
            820                     825                     830
Trp Ala Thr Leu Ala Gln Lys Leu Gly Leu Gly Ile Leu Asn Asn Ala
        835                     840                     845
Phe Arg Leu Ser Pro Ala Pro Ser Lys Thr Leu Met Asp Asn Tyr Glu
    850                     855                     860
Val Ser Gly Gly Thr Val Arg Glu Leu Val Glu Ala Leu Arg Gln Met
865                     870                     875                     880
Gly Tyr Thr Glu Ala Ile Glu Val Ile Gln Ala Ala Ser Ser Pro Val
                885                     890                     895
Lys Thr Thr Ser Gln Ala His Ser Leu Pro Leu Ser Pro Ala Ser Thr
            900                     905                     910
Arg Gln Gln Ile Asp Glu Leu Arg Asp Ser Asp Ser Val Cys Asp Thr
            915                     920                     925
Gly Val Glu Thr Ser Phe Arg Lys Leu Ser Phe Thr Glu Ser Leu Thr
    930                     935                     940
Ser Gly Ala Ser Leu Leu Thr Leu Asn Lys Met Pro His Asp Tyr Gly
945                     950                     955                     960
Gln Glu Gly Pro Leu Glu Gly Lys Ile
                965


<210> 40
<211> 96
<212> DNA
<213> Artificial Sequence


<220>
<223> engineered regulatory sequence


<400> 40
gctagccccg ccccgttgac gcaaatgggc ggtaggcgtg tacggtggga ggtttatata      60
agcagagctc gtttagtgaa ccgtcagatc agatct                                96


<210> 41
<211> 154
<212> DNA
<213> Artificial Sequence


<220>
<223> engineered regulatory sequence


<400> 41
gctagcgccc aaattgggat tttccaaaaa gccgaaattg ggattttcca aaaaccgccg      60
atcgcccgcc ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg tttatataag     120
cagagctcgt ttagtgaacc gtcagatcag atct                                154


<210> 42
<211> 212
<212> DNA
<213> Artificial Sequence


<220>
<223> engineered regulatory sequence


<400> 42
acgcgtgccc aaattgggat tttccaaaaa gccgaaattg ggattttcca aaaaccgcgc      60
```

13

```
tagcgcccaa attgggattt tccaaaaagc cgaaattggg attttccaaa aaccgccgat      120
cgcccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtt tatataagca      180
gagctcgttt agtgaaccgt cagatcagat ct                                    212


<210> 43
<211> 96
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 43
gctagccccg ccccgttgac gcaaatgggc ggtaggcgtg tacggtggga ggtctatata       60
agcagagctc gtttagtgaa ccgtcagatc agatct                                 96


<210> 44
<211> 154
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 44
gctagcgccc aggtcgggat tttccgagga gccgaggtcg ggattttccg aggaccgccg       60
atcgcccgcc ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg cctatataag      120
cagagctcgt ttagtgaacc gtcagatcag atct                                  154


<210> 45
<211> 154
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 45
gctagcgccc aggtcgggat tttccgagga gccgaggtcg ggattttccg aggaccgccg       60
atcgcccgcc ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg cctatataag      120
cagagctcgt ttagtgaacc gtcagatcag atct                                  154


<210> 46
<211> 762
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered promoter construct

<400> 46
ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaattaatat       60
tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc      120
atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat      180
tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa      240
tggcccgcct ggctgaccgc ccaacgaccc cgcccattg acgtcaataa tgacgtatgt      300
tcccatagta acgcaaatag gattttccca ttaacgtcaa tgggtggagt atttacggta      360
```

14

```
aactgcccac ttggcagtac atcaagtgta tcatatgcca agtccgcccc ctattgacgt        420
caatgacggt aaatggcccg cctggcatta tgcccagtac atgactttat gggattttcc        480
tatttggcag tacatctacg tattagtcat cgctattacc atggtgatgc ggttttggca        540
gtacaccaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat        600
tgacgtcaat gggagtttgt tttggcacca aggtaaaagg gattttccaa aatgtcgtaa        660
caactgcgat cgcccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtt        720
tatataagca gagctcgttt agtgaaccgt cagatcaagc tt                           762


<210> 47
<211> 762
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered promoter construct

<400> 47
ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaattaatat         60
tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc        120
atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat        180
tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa        240
tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt        300
tcccatagta acgcaaatat ccccgggaaa ttaacgtcaa tgggtggagt atttacggta        360
aactgcccac ttggcagtac atcaagtgta tcatatgcca agtccgcccc ctattgacgt        420
caatgacggt aaatggcccg cctggcatta tgcccagtac atgactttat tctcgaggaa        480
tatttggcag tacatctacg tattagtcat cgctattacc atggtgatgc ggttttggca        540
gtacaccaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat        600
tgacgtcaat gggagtttgt tttggcacca aggtaaaatt acgcgtaaaa aatgtcgtaa        660
caactgcgat cgcccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtt        720
gctagccgca gagctcgttt agtgaaccgt cagatcaagc tt                           762


<210> 48
<211> 762
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered promoter construct

<400> 48
ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaatcaatat         60
tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc        120
atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat        180
tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa        240
tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt        300
tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta        360
aactgcccac ttggcagtac atcaagtgta tcatatgcca agtccgcccc ctattgacgt        420
caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttac gggactttcc        480
tacttggcag tacatctacg tattagtcat cgctattacc atggtgatgc ggttttggca        540
gtacaccaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat        600
tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa        660
caactgcgat cgcccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtc        720
tatataagca gagctcgttt agtgaaccgt cagatcaagc tt                           762


<210> 49
<211> 12
<212> DNA
```

15

<213> Artificial Sequence

<220>
<223> wild type regulatory sequence

<400> 49
gactgtttgt tt   12

<210> 50
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> wild type regulatory sequence

<400> 50
aggactcttg ga   12

<210> 51
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> wild type regulatory sequence

<400> 51
tactaggagg ctgtaggcat aaattggtct gcgcaccagc accatg   46

<210> 52
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 52
tactaggagg ctgtaggcat aaattagtct gcgcaccagc accatg   46

<210> 53
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 53
tactaggatt agtgcttaag cccttggtct gcgcaccagc accatg   46

<210> 54
<211> 46
<212> DNA
<213> Artificial Sequence

```
<220>
<223> engineered regulatory sequence

<400> 54
tactaggagg ctgtaggcat aaagctcgag tatacaacgc accatg                        46

<210> 55
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 55
tactaggagg ctgtaggcat aaatgcgtaa aagcaccagc accatgcaac                    50

<210> 56
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 56
tactaggagg ctgtaggcat aaattaaaaa acgcaccagc accatgcaac                    50

<210> 57
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 57
tactaggagg ctgtaggcat aaattaatcc gcgcaccagc accatgcaac                    50

<210> 58
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 58
accttgaggc atacttcaaa gactgttgat ttagcgaata agaggagttg g                  51

<210> 59
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence
```

<400> 59
accttgaggc atacttcaaa gactgtttat tttaataacg ggaggagttg g   51

<210> 60
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> engineered regulatory sequence

<400> 60
accttgaggc atacttcaaa gactgtttat ttaaggactg ggaggagttg g   51

<210> 61
<211> 6513
<212> DNA
<213> Artificial Sequence

<220>
<223> heterologous nucleic acid construct

<400> 61
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta   60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc   120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg   180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc   240
gcctggctga ccgcccaacg accccgccc attgacgtca ataatgacgt atgttcccat   300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc   360
ccacttggca gtacatcaag tgtatcatat gccaagtccg cccctattg acgtcaatga   420
cggtaaatgg cccgcctggc attatgccca gtacatgacc ttacgggact ttcctacttg   480
gcagtacatc tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacac   540
caatgggcgt ggatagcggt ttgactcacg gggatttcca gtctccacc ccattgacgt   600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactg   660
cgatcgcccg ccccgttgac gcaaatgggc ggtaggcgtg tacggtggga ggtctatata   720
agcagagctc gtttagtgaa ccgtcagatc actagaagct ttattgcggt agtttatcac   780
agttaaattg ctaacgcagt cagtgcttct gacacaacag tctcgaactt aagctgcagt   840
gactctctta aggtagcctt gcagaagttg gtcgtgaggc actgggcagg taagtatcaa   900
ggttacaaga caggtttaag gagaccaata gaaactgggc ttgtcgagac agagaagact   960
cttgcgtttc tgataggcac ctattggtct tactgacatc cactttgcct ttctctccac   1020
aggtgtccac tcccagttca attacagctc ttaaggctag agtacttaat acgactcact   1080
ataggctagc cagcttgaag caagcctcct gaaagatgga ggcgtcgctg ccggcccagg   1140
ccgccgagac ggaggaggtg ggtcttttcg tcgaaaaata cctccggtcc gatgtcgcgc   1200
cggcggaaat tgtcgcgctc atgcgcaacc tcaacagcct gatgggacgc acgcggttta   1260
tttacctggc gttgctggag gcctgtctcc gcgttccat ggccacccgc agcagcgcca   1320
tatttcggcg gatctatgac cactacgcca cgggcgtcat ccccacgatc aacgtcaccg   1380
gagagctgga gctcgtggcc ctgccccca ccctgaacgt aaccccgtc tgggagctgt   1440
tgtgcctgtg cagcaccatg gccgcgcgcc tgcattggga ctcggcggcc gggggatctg   1500
ggaggacctt cggccccgat gacgtgctgg acctactgac cccccactac gaccgctaca   1560
tgcagctggt gttcgaactg ggccactgta cgtaaccga cggacttctg ctctcggagg   1620
aagccgtcaa gcgcgtcgcc gacgccctaa gcggctgtcc cccgcgcggg tccgttagcg   1680
agacggacca cgcggtggcg ctgttcaaga taatctgggg cgaactgttt ggcgtgcaga   1740
tggccaaaag cacgcagacg tttcccgggg cggggcgcgt taaaaacctc accaaacaga   1800
caatcgtggg gttgttggac gcccaccaca tcgaccacag cgcctgccgg acccacaggc   1860
agctgtacgc cctgcttatg gcccacaagc gggagtttgc gggcgcgcgc ttcaagctac   1920
gcgtgcccgc gtggggggcgc tgtttgcgca cgcactcatc cagcgccaac cccaacgctg   1980

```
acatcatcct ggaggcggcg ctgtcggagc tccccaccga ggcctggccc atgatgcagg   2040
gggcggtgaa ctttagcacc ctaatgaagc tactgtcttc tatcgaacaa gcatgcccaa   2100
aaaagaagag aaaggtagat gaattcccgg ggatctcgac ggcccccccg accgatgtca   2160
gcctggggga cgagctccac ttagacggcg aggacgtggc gatggcgcat gccgacgcgc   2220
tagacgattt cgatctggac atgttggggg acggggattc cccgggtccg ggatcgccag   2280
ggatccgtcg acttgacgcg ttgatatcat ctagagcggc cgcaggtacc tgaataacta   2340
aggccgcttc cctttagtga gggttaatgc ttcgagcaga catgataaga tacattgatg   2400
agtttggaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg   2460
atgctattgc tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt   2520
gcattcattt tatgtttcag gttcagggggg agatgtggga ggttttttaa agcaagtaaa   2580
acctctacaa atgtggtaaa atccgataag gatcgattcc ggagcctgaa tggcgaatgg   2640
acgcgccctg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc acgtgaccgc   2700
tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct ttctcgccac   2760
gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt tccgatttag   2820
tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttcac gtagtgggcc   2880
atcgccctga tagacggttt ttcgcccttt gacgttggag tccacgttct ttaatagtgg   2940
actcttgttc caaactggaa caacactcaa ccctatctcg gtctattctt ttgatttata   3000
agggattttg ccgatttcgg cctattggtt aaaaaatgag ctgatttaac aaaaatttaa   3060
cgcgaatttt aacaaaatat taacgcttac aatttcgcct gtgtaccttc tgaggcggaa   3120
agaaccagct gtggaatgtg tgtcagttag ggtgtggaaa gtccccaggc tccccagcag   3180
gcagaagtat gcaaagcatg catctcaatt agtcagcaac caggtgtgga aagtccccag   3240
gctccccagc aggcagaagt atgcaaagca tgcatctcaa ttagtcagca accatagtcc   3300
cgcccctaac tccgcccatc ccgcccctaa ctccgcccag ttccgcccat tctccgcccc   3360
atggctgact aattttttt atttatgcag aggccgaggc cgcctcggcc tctgagctat   3420
tccagaagta gtgaggaggc ttttttggag gcctaggctt ttgcaaaaag cttgattctt   3480
ctgacacaac agtctcgaac ttaaggctag agccaccatg attgaacaag atggattgca   3540
cgcaggttct ccggccgctt gggtggagag gctattcggc tatgactggg cacaacagac   3600
aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc cggttctttt   3660
tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc   3720
gtggctggcc acgacgggcg ttccttgcgc agctgtgctc gacgttgtca ctgaagcggg   3780
aagggactgg ctgctattgg gcgaagtgcc ggggcaggat ctcctgtcat ctcaccttgc   3840
tcctgccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgcttgatcc   3900
ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac gtactcggat   3960
ggaagccggt cttgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc   4020
cgaactgttc gccaggctca aggcgcgcat gcccgacggc gaggatctcg tcgtgaccca   4080
tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgcttttctg gattcatcga   4140
ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggcta cccgtgatat   4200
tgctgaagag cttggcggcg aatgggctga ccgcttcctc gtgctttacg gtatcgccgc   4260
tcccgattcg cagcgcatcg ccttctatcg ccttcttgac gagttcttct gagcgggact   4320
ctggggttcg aaatgaccga ccaagcgacg cccaacctgc catcacgatg ccgcaataa   4380
aatatcttta ttttcattac atctgtgtgt tggttttttg tgtgaagatc cgcgtatggt   4440
gcactctcag tacaatctgc tctgatgccg catagttaag ccagccccga cacccgccaa   4500
caccgctga cgcgccctga cgggcttgtc tgctcccggc atccgcttac agacaagctg   4560
tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg aaacgcgcga   4620
gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt   4680
cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt   4740
tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat   4800
aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt   4860
ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg   4920
ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga   4980
tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc   5040
tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac   5100
actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg   5160
gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca   5220
acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg   5280
gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg   5340
acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg   5400
```

```
gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag    5460
ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg    5520
gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct    5580
cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac    5640
agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac caagtttact    5700
catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga    5760
tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt    5820
cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct    5880
gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc    5940
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc    6000
ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc    6060
tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg    6120
ggttggactc aagacgatag ttaccggata aggcgcagcg tcgggctga acggggggtt    6180
cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg    6240
agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg    6300
gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt    6360
atagtcctgt cggtttcgc cacctctgac ttgagcgtcg attttttgtga tgctcgtcag    6420
ggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt    6480
gctggccttt tgctcacatg gctcgacaga tct                                  6513

<210> 62
<211> 6439
<212> DNA
<213> Artificial Sequence

<220>
<223> heterologous nucleic acid construct

<400> 62
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta     60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc    120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg    180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc    240
gcctggctga ccgcccaacg accccgcccc attgacgtca ataatgacgt atgttcccat    300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc    360
ccacttggca gtacatcaag tgtatcatat gccaagtccg cccccattg acgtcaatga    420
cggtaaatgg cccgcctggc attatgccca gtacatgacc ttacgggact ttcctacttg    480
gcagtacatc tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacac    540
caatgggcgt ggatagcggt ttgactcacg gggatttcca agtctccacc ccattgacgt    600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactg    660
cgatcgcccg ccccgttgac gcaaatgggc ggtaggcgtg tacggtggga ggtctatata    720
agcagagctc gtttagtgaa ccgtcagatc actagaagct ttattgcggt agtttatcac    780
agttaaattg ctaacgcagt cagtgcttct gacacaacag tctcgaactt aagctgcagt    840
gactctctta aggtagcctt gcagaagttg gtcgtgaggc actgggcagg taagtatcaa    900
ggttacaaga caggtttaag gagaccaata gaaactgggc ttgtcgagac agagaagact    960
cttgcgtttc tgataggcac ctattggtct tactgacatc cactttgcct ttctctccac   1020
aggtgtccac tcccagttca attacagctc ttaaggctag agtacttaat acgactcact   1080
ataggctagc cagcttgaag caagcctcct gaaagatgga ggcgtcgctg ccggcccagg   1140
ccgccgagac ggaggaggtg ggtctttttcg tcgaaaaata cctccggtcc gatgtcgcgc   1200
cggcggaaat tgtcgcgctc atgcgcaacc tcaacagcct gatgggacgc acgcggttta   1260
tttacctggc gttgctggag cctgtctcc gcgttcccat ggccacccgc agcagcgcca   1320
tatttcggcg gatctatgac cactacgcca cgggcgtcat ccccacgatc aacgtcaccg   1380
gagagctgga gctcgtggcc ctgcccccca ccctgaacgt aaccccgtc tgggagctgt   1440
tgtgcctgtg cagcaccatg gccgcgcgcc tgcattggga ctcggcggcc gggggatctg   1500
ggaggacctt cggccccgat gacgtgctgg acctactgac cccccactac gaccgctaca   1560
tgcagctggt gttcgaactg ggccactgta acgtaaccga cggacttctg ctctcggagg   1620
aagccgtcaa gcgcgtcgcc gacgccctaa gcggctgtcc cccgcgcggg tccgttagcg   1680

                                   20
```

```
agacggacca cgcggtggcg ctgttcaaga taatctgggg cgaactgttt ggcgtgcaga    1740
tggccaaaag cacgcagacg tttcccgggg cggggcgcgt taaaaacctc accaaacaga    1800
caatcgtggg gttgttggac gcccaccaca tcgaccacag cgcctgccgg acccacaggc    1860
agctgtacgc cctgcttatg gcccacaagc gggagtttgc gggcgcgcgc ttcaagctac    1920
gcgtgcccgc gtgggggcgc tgtttgcgca cgcactcatc cagcgccaac cccaacgctg    1980
acatcatcct ggaggcggcg ctgtcggagc tccccaccga ggcctggccc atgatgcagg    2040
gggcggtgaa ctttagcacc ctaccaaaaa agaagagaaa ggtagatcgg acactggtga    2100
ccttcaagga tgtatttgtg dacttcacca gggaggagtg gaagctgctg gacactgctc    2160
agcagatcgt gtacagaaat gtgatgctgg agaactataa gaacctggtt tccttgggtt    2220
attgatgaga tatcatctag agcggccgca ggtacctgaa taactaaggc cgcttccctt    2280
tagtgagggt taatgcttcg agcagacatg ataagataca ttgatgagtt tggacaaacc    2340
acaactagaa tgcagtgaaa aaaatgcttt atttgtgaaa tttgtgatgc tattgcttta    2400
tttgtaacca ttataagctg caataaacaa gttaacaaca caattgcat tcattttatg    2460
tttcaggttc agggggagat gtgggaggtt ttttaaagca agtaaaacct ctacaaatgt    2520
ggtaaaatcc gataaggatc gattccggag cctgaatggc gaatggacgc gccctgtagc    2580
ggcgcattaa gcgcggcggg tgtggtggtt acgcgcacgt gaccgctaca cttgccagcg    2640
ccctagcgcc cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc    2700
cccgtcaagc tctaaatcgg gggctccctt tagggttccg atttagtgct ttacggcacc    2760
tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg ccctgataga    2820
cggttttttcg cccttttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa    2880
ctggaacaac actcaaccct atctcggtct attcttttga tttataaggg attttgccga    2940
tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattttaaca    3000
aaatattaac gcttacaatt tcgcctgtgt accttctgag gcggaaagaa ccagctgtgg    3060
aatgtgtgtc agttaggtg tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa    3120
agcatgcatc tcaattagtc agcaaccagg tgtggaaagt ccccaggctc cccagcaggc    3180
agaagtatgc aaagcatgca tctcaattag tcagcaacca tagtcccgcc cctaactccg    3240
cccatcccgc ccctaactcc gcccagttcc gcccattctc cgccccatgg ctgactaatt    3300
ttttttattt atgcagaggc cgaggccgcc tcggcctctg agctattcca gaagtagtga    3360
ggaggctttt ttggaggcct aggcttttgc aaaaagcttg attcttctga cacaacagtc    3420
tcgaacttaa ggctagagcc accatgattg aacaagatgg attgcacgca ggttctccgg    3480
ccgcttgggt ggagaggcta ttcggctatg actgggcaca acagacaatc ggctgctctg    3540
atgccgccgt gttccggctg tcagcgcagg ggcgcccggt tctttttgtc aagaccgacc    3600
tgtccggtgc cctgaatgaa ctgcaggacg aggcagcgcg gctatcgtgg ctggccacga    3660
cgggcgttcc ttgcgcagct gtgctcgacg ttgtcactga agcgggaagg gactggctgc    3720
tattgggcga agtgccgggg caggatctcc tgtcatctca ccttgctcct gccgagaaag    3780
tatccatcat ggctgatgca atgcggcggc tgcatacgct tgatccggct acctgcccat    3840
tcgaccacca agcgaaacat cgcatcgagc gagcacgtac tcggatggaa gccggtcttg    3900
tcgatcagga tgatctggac gaagagcatc aggggctcgc gccagccgaa ctgttcgcca    3960
ggctcaaggc gcgcatgccc gacggcgagg atctcgtcgt gacccatggc gatgcctgct    4020
tgccgaatat catggtggaa aatggccgct tttctggatt catcgactgt ggccggctgg    4080
gtgtggcgga ccgctatcag gacatagcgt tggctacccg tgatattgct gaagagcttg    4140
gcggcgaatg ggctgaccgc ttcctcgtgc tttacggtat cgccgctccc gattcgcagc    4200
gcatcgcctt ctatcgcctt cttgacgagt tcttctgagc gggactctgg ggttcgaaat    4260
gaccgaccaa gcgacgccca acctgccatc acgatggccg caataaaata tctttatttt    4320
cattacatct gtgtgttggt tttttgtgtg aagatccgcg tatggtgcac tctcagtaca    4380
atctgctctg atgccgcata gttaagccag ccccgacacc cgccaacacc cgctgacgcg    4440
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg    4500
agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgagacg aaagggcctc    4560
gtgatacgcc tatttttata ggttaatgtc atgataataa tggtttctta gacgtcaggt    4620
ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tattttttcta aatacattca    4680
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg    4740
aagagtatga gtattcaaca tttccgtgtc gcccttattc ccttttttgc ggcattttgc    4800
cttcctgttt ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg    4860
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt    4920
cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta    4980
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat    5040
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga    5100
```

21

```
gaattatgca gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca      5160
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact      5220
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc      5280
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact      5340
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt      5400
ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt      5460
gggtctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt      5520
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata      5580
ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag      5640
attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat      5700
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa      5760
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca      5820
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt      5880
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg      5940
tagttaggcc accacttcaa gaactctgta gcaccgccta cacctcgc tctgctaatc      6000
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga      6060
cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc      6120
agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc      6180
gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca      6240
ggagagcgca cgagggagct tccagggggga aacgcctggt atctttatag tcctgtcggg      6300
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcagggg gcggagccta      6360
tggaaaaacg ccagcaacgc ggcctttttta cggttcctgg cctttttgctg gccttttgct      6420
cacatggctc gacagatct                                                 6439
```